Available online at www.sciencedirect.com
ScienceDirect
Molecular Phylogenetics and Evolution 43 (2007) 283–297
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

A molecular phylogeny of *Anopheles annulipes* (Diptera: Culicidae) sensu lato: The most species-rich anopheline complex

D.H. Foley [a,*], R.C. Wilkerson [a], R.D. Cooper [b], M.E. Volovsek [a], J.H. Bryan [c]

[a] *Department of Entomology, Walter Reed Army Institute of Research, 503 Robert Grant Avenue, Silver Spring, MD 20910, USA*
[b] *Australian Army Malaria Institute, Gallipoli Barracks, Enoggera, Queensland 4051, Australia*
[c] *Tropical Health Program and Department of Zoology and Entomology, University of Queensland, Brisbane 4072, Australia*

Received 29 March 2006; revised 27 September 2006; accepted 9 October 2006
Available online 17 October 2006

Abstract

The Australasian Annulipes Complex is the most species-rich among *Anopheles* mosquitoes, with at least 15 sibling species suspected. Members of this complex are the most likely vectors of malaria in the past in southern Australia and are involved in the spread of myxomatosis among rabbits. In this, the first comprehensive molecular study of the Annulipes Complex, 23 ITS2 rDNA variants were detected from collections throughout Australia and Papua New Guinea, including diagnostic variants for the previously identified *An. annulipes* species A–G. Specimens of each ITS2 variant were sequenced for portions of the mitochondrial COI, COII and nuclear EF-1α genes. Partitioned Bayesian and Maximum Parsimony analyses confirmed the monophyly of the Annulipes Complex and revealed at least 17 clades that we designate species A–Q. These species belong to two major clades, one in the north and one mainly in the south, suggesting that climate was a driver of species radiation. We found that 65% (11) of the 17 sibling species recorded here had unique COI sequences, suggesting that DNA barcoding will be useful for diagnosing species within the Annulipes Complex. A comparison of the taxa revealed morphological characters that may be diagnostic for some species. Our results substantially increase the size of the subgenus *Cellia* in Australasia, and will assist species-level studies of the Annulipes Complex.

Keywords: *Anopheles annulipes*; Phylogenetics; Bayesian; Maximum parsimony; Sibling species; Species complex; Myxomatosis; DNA barcoding; Australia; Papua New Guinea; ITS2; COI; COII; EF-1α

1. Introduction

Anopheles annulipes sensu lato (s.l.) Walker is the most ubiquitous anopheline mosquito in Australia and also occurs in New Guinea (NG) (Lee et al., 1987). This taxon has been implicated in past malaria outbreaks in the north and northwest of Australia and is the presumed vector of the occasional case of indigenous malaria in southern areas (Black, 1972). A number of arboviruses have been recovered from *An. annulipes* s.l. (Russell, 1995). This taxon is also the most important vector of myxoma virus among rabbits in many areas of Australia (Fenner and Ratcliffe, 1965; Parer and Korn, 1989).

Frank Walker described *Anopheles annulipes* in 1856 from a collection from Van Diemen's Land (Tasmania = TAS). This taxon exhibits extensive morphological variation that has resulted in various taxonomic interpretations; five names have been synonymized under *An. annulipes* (*An. musivus* Skuse 1889, *An. mastersi* Skuse 1889, *An. perplexus* Taylor, 1943, *An. perplexus* var. *persimilis* Taylor, 1943 and *An. derricki* Taylor, 1943). The possibility of more than one biological form of *An. annulipes* was proposed as an explanation for geographical differences in the ability of myxomatosis to control rabbits (Fenner and Ratcliffe, 1965).

Chris Green used data from cross-matings and the banding pattern of polytene chromosomes to demonstrate that *An. annulipes* in Australia comprises at least four species

* Corresponding author. Present address: Walter Reed Biosystematics Unit, Smithsonian Institution MSC MRC534, 4210 Silver Hill Road, Suitland, MD 20746, USA. Fax: +1 301 238 3168.
E-mail address: foleydes@si.edu (D.H. Foley).

1055-7903/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2006.10.008

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
27 SEP 2006		**00-00-2006 to 00-00-2006**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
A molecular phylogeny of Anopheles annulipes (Diptera: Culicidae) sensu lato: The most species-rich anopheline complex	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Walter Reed Army Institute of Research,Department of Entomology,503 Robert Grant Avenue,Silver Spring,MD,20910**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**15**	

(species A–D) (unpublished; Green, 1972; Paterson and James, 1973). Using the same methodology, Booth and Bryan (1986) named three new species (species E–G) from locations in eastern Australia and suggested that additional species occur in the Northern Territory (NT) and Queensland (QLD). They speculated that at least 10 sibling species make up this taxon in Australia and Booth et al. (1987) provided a map showing the distribution of the sibling species. Intensive studies based on allozyme markers of species A and G at Griffith in New South Wales (NSW) showed that they do not interbreed in nature, and have significant differences in phenology, behavior and ecology (Foley and Bryan, 1991a,b; Bryan et al., 1991; Foley et al., 1992). Putative representatives of species A, B, D, and G were included in a phylogeny of Australasian anophelines using sequence data for cytochrome *c* oxidase subunit II (COII); the Annulipes Complex appeared to be monophyletic with strong branch support (Foley et al., 1998). Recently, an allozyme survey of the Annulipes Complex from 61 sites around Australia indicated 15–25 putative species (Foley et al., in press). Additional species may be found in Papua New Guinea (PNG) based on the results of distribution modelling (Foley et al., in press) and the identification of two distinct ribosomal DNA (rDNA) internal transcribed spacer subunit 2 (ITS2) RFLP genotypes (Cooper et al., in press). The molecular delineation of species boundaries is fundamental to any further study of this taxon, and past studies will have to be reassessed in the light of its multispecies status.

In this, the first comprehensive molecular study of the Annulipes Complex, we aimed to test the monophyly of the group, to determine the number of species present, and to provide molecular data for species diagnosis. Nuclear and mitochondrial genes have different strengths and weaknesses when it comes to phylogenetic reconstruction (e.g. Lin and Danforth, 2004). We used portions of two nuclear genes (nDNA) and two mitochondrial genes (mtDNA), representing coding and non-coding regions, in maximum parsimony and partitioned Bayesian phylogenetic analyses. A partitioned Bayesian (i.e. mixed model) approach allows the simultaneous analysis of different genes and codon positions, incorporating realistic models of their DNA sequence evolution, which is expected to reduce systematic error and increase phylogenetic signal, even in cases of rapid radiation (Brandley et al., 2005). We included representatives of many of the chromosome- and allozyme-defined species and forms from previous studies, and specimens collected throughout the range of *An. annulipes*, including from many of the type localities.

2. Materials and methods

2.1. Mosquito collections and identification

Adult females of *An. annulipes* and outgroup taxa were identified using the keys of Lee et al. (1987). Details of the collection and identification of specimens used in this study are given in Table 1. Some specimens used in this project had been identified according to polytene chromosome banding patterns or were the same specimens or F_1 progeny of specimens used in Booth and Bryan (1986) and Foley and Bryan (1991a,b). These specimens (according to species, location, code [specimen number]) were: species A (Mildura: Army [56]; Griffith: Misc 10 [61], Misc 15 [36], 3-F_1 [39]), species C (Homebush: D. Booth 14/4 [259], Termeil: 28F_1 [38]), species E (Woronora: D. Booth 390 [129]), species F (Dunoon: D. Booth 421 [127], D. Booth 421 [260]), species G (Griffith: D. Booth 9 [261], 13 [262], P10(6) [62], P18(8) [40]), and the type 1 Mataranka form of Booth and Bryan (1986) (Mataranka D. Booth Mat IX [128]). In addition, species C included specimens from the 'Standard' strain colony; derived from near Sydney in 1972 and maintained at the Army Malaria Research Unit, Ingleburn [258]. Morphology, species distribution data (Green, unpublished; Liehne, 1991) and allozyme electrophoresis (unpublished) were used to tentatively identify *An. annulipes* species B and D. Specimens from PNG included specimens from highland and lowland sites that Cooper et al. (in press) found differed in ITS2 RFLP genotypes. Specimens of *Anopheles farauti* Laveran, *An. amictus* Edwards, *An. hilli* Lee and Woodhill, *An. novaguinensis* Venhuis, and *An. meraukensis* Venhuis, representing most of the currently known species of the subgenus *Cellia* from Australia, were included in the phylogenetic analysis (see Table 1 for collection details).

2.2. DNA extraction

Specimens had been stored dry at −70 to −80 °C for up to 20 years and were transferred to 100% ethanol at −80 °C for up to one year prior to DNA extraction. Often only sequences less than 400 bp could be reliably amplified, which affected the choice of DNA extraction technique and required the design of internal primers (see below). Initially, DNA from single legs was extracted but legs were hard to grind to completion and PCR results were not consistent, therefore mosquito abdomens were used. Within pools of mosquitoes, abdomens had the additional advantage that they could be unequivocally associated with individual mosquitoes. DNA was extracted with DNAzol (Invitrogen, Carlsbad, CA, USA), a guanidine thiocyanate-detergent lysing solution, using a modification of the manufacturer's protocol for tissues. Guanidine thiocyanate gives good amplification of lower quality DNA (e.g. Rohland et al., 2004). Abdomens were briefly air-dried then placed in a 1.5 ml microcentrifuge tube containing 125 µl of DNAzol reagent. Abdomens were individually ground in 125 µl of DNAzol reagent with a pellet pestle (Kimble/Kontes, Vineland, NJ, USA), then a further 125 µl of reagent was added, and the contents vortexed and stored at room temperature for 5 min. The tubes were then centrifuged for 10 min at 10,000*g* to pellet insoluble tissue fragments and RNA. The supernatant was transferred to a new tube and 125 µl of 100% ethanol was added to precipitate the DNA. Samples were mixed by inverting the tubes several times and then

Table 1
Collection details of *Anopheles annulipes* s.l. from Australia and Papua New Guinea

Locality number and name[a]	Longitude	Latitude	Species (*n*)	Specimen	GenBank accession numbers			
					ITS2	EF-1α	COI	COII
Outgroup taxa								
Katherine, NT	132°16′E	14°28′S	*An. amictus*	(35)	—	DQ420553	DQ420425	DQ420489
Seram Is, Maluku, Indonesia (Kufpec Camp I, V-2002 WKI Sk)[d]	130°26′E	3°10′S	*An. farauti*	(110)	—	DQ420552	DQ420424	DQ420488
Darwin, NT	130°50′E	12°27′S	*An. hilli*	(41)	—	DQ420554	DQ420426	DQ420490
Darwin, NT	130°50′E	12°27′S	*An. meraukensis*	(11)	—	DQ420556	DQ420428	DQ420492
Irvinebank, QLD	145°12′E	17°26′S	*An. novaguinensis*	(42)	—	DQ420555	DQ420427	DQ420491
Annulipes Complex								
1 Basalt R.[b]	145°46′E	19°37′S	H-11,I-12,Q-21[2]	H-11(132)	DQ420388	DQ420580	DQ420452	DQ420516
				Q-21(136)	DQ420420	DQ420612	DQ420484	DQ420548
				Q-21(153)	DQ420421	DQ420613	DQ420485	DQ420549
2 Batavia	142°40′E	12°40′S	K-14[3]					
3 Burdekin R.	147°10′E	19°45′S	D-7[6],H-11	H-11(135)	DQ420389	DQ420581	DQ420453	DQ420517
4 Clarke R.[b]	145°26′E	19°13′S	I-12[3],J-13, K-14,O-19	I-12(131)	DQ420391	DQ420583	DQ420455	DQ420519
				O-19(154)	DQ420406	DQ420598	DQ420470	DQ420534
5 Dilulu[b]	150°16′E	23°53′S	?-21[3]	21(159)	DQ420422	DQ420614	DQ420486	DQ420550
6 Eidsvold[b]	151°07′E	25°22′S	E-8[3]					
7 Eungella[c]	148°30′E	21°08′S	E-8,F-9[3],Q-21	E-8(44)	DQ420376	DQ420568	DQ420440	DQ420504
8 Gladstone[b]	151°04′E	23°54′S	K-14					
9 Horn Is[b, c]	142°17′E	10°35′S	K-14[11]	K-14(57)	DQ420395	DQ420587	DQ420459	DQ420523
10 Innot Hot Springs[b]	145°14′E	17°40′S	K-14[2]	K-14(76)	DQ420397	DQ420589	DQ420461	DQ420525
11 Irvinebank[b]	145°12′E	17°26′S	F-9,K-14[4],O-19	F-9(87)	DQ420380	DQ420572	DQ420444	DQ420508
				O-19(248)	DQ420407	DQ420599	DQ420471	DQ420535
				O-19(256)	DQ420408	DQ420600	DQ420472	DQ420536
12 Kennedy Ck[c]	144°26′E	15°43′S	K-14[3]					
13 Kowanyama	141°45′E	15°28′S	D-7[2],J-13[2]	D-7(100)	DQ420375	DQ420567	DQ420439	DQ420503
				J-13(60)	DQ420393	DQ420585	DQ420457	DQ420521
				J-13(99)	DQ420394	DQ420586	DQ420458	DQ420522
14 Lake Manchester[b]	152°45′E	27°28′S	E-8[4]					
15 Prince of Wales Is[b]	142°07′E	10°43′S	K-14[2]					
16 Seisia[c]	142°22′E	10°51′S	K-14[6]					
17 Townsville region[b,c]	146°45′E	19°26′S	D-7[8],I-12[3], O-19[2],?-23	D-7(97)	DQ420374	DQ420566	DQ420438	DQ420502
				I-12(83)	DQ420390	DQ420582	DQ420454	DQ420518
				23(67)	DQ420416	DQ420608	DQ420480	DQ420544
18 Umagico[c]	142°23′E	10°53′S	K-14					
19 Bateman's Bay[b]	150°15′E	35°44′S	C-5	C-5(144)	DQ420386	DQ420578	DQ420450	DQ420514
20 Castle Hill Colony [Standard]	151°00′E	33°43′S	C-4					
21 Coffs Harbour[d]	153°08′E	30°18′S	C-5	C-5(130)	DQ420385	DQ420577	DQ420449	DQ420513
22 Condobolin[c]	147°09′E	33°05′S	A-1[2]					
23 Conjola[c]	150°26′E	35°13′S	C-4[2],C-5	C-5(108)	DQ420409	DQ420601	DQ420473	DQ420537
24 Dunoon[b]	153°19′E	28°41′S	F-9[3]	F-9(48)	DQ420379	DQ420571	DQ420443	DQ420507
				F-9(127)	DQ420381	DQ420573	DQ420445	DQ420509
25 Forbes[c]	148°01′E	33°23′S	A-1[3],G-10					
26 Griffith[c]	146°02′E	34°17′S	A-1[4],G-10[4]	A-1(39)	DQ420367	DQ420559	DQ420431	DQ420495
				G-10(40)	DQ420382	DQ420574	DQ420446	DQ420510
27 Homebush	151°05′E	33°52′S	C-4					
28 Lord Howe Is[b]	159°05′E	31°33′S	C-4,C-5,C-6	C-4(49)	DQ420372	DQ420564	DQ420436	DQ420500
				C-5(206)	DQ420410	DQ420602	DQ420474	DQ420538
				C-6(207)	DQ420418	DQ420610	DQ420482	DQ420546
29 McCarrs Ck[b]	151°16′E	33°40′S	A-1,E-8					
34 Mittagong[b]	150°27′E	34°27′S	E-8[3]					
30 Oberon [Fish R.][d]	149°51′E	33°42′S	G-10	G-10(255)	DQ420384	DQ420576	DQ420448	DQ420512
31 Tenterfield [Reedy Ck][b]	151°50′E	29°09′S	E-8[2],H-11[3]	H-11(32)	DQ420387	DQ420579	DQ420451	DQ420515
32 Termeil State Forest[c]	150°22′E	35°26′S	C-4[2]	C-4(38)	DQ420371	DQ420563	DQ420435	DQ420499
33 Walgett[c]	148°07′E	30°01′S	G-10[2]					
34 Warren[c]	147°50′E	31°42′S	A-1					
35 Woronora[b]	151°02′E	34°02′S	E-8[3],H-11	E-8(129)	DQ420378	DQ420570	DQ420442	DQ420506
36 Busselton [BSN38,39][c]	115°21′E	33°39′S	A-1[4]	A-1(29)	DQ420366	DQ420558	DQ420430	DQ420494
37 Capel Shire [BSN62-64,80,81][c]	115°33′E	33°33′S	A-1[15]					
38 Kununurra [10544][c]	128°44′E	15°46′S	D-7[7]	D-7(68)	DQ420373	DQ420565	DQ420437	DQ420501

(*continued on next page*)

Table 1 (*continued*)

Locality number and name[a]		Longitude	Latitude	Species (*n*)	Specimen	GenBank accession numbers			
						ITS2	EF-1α	COI	COII
39	Minnie R. Derby Shire [91WK24,25][c]	123°36′E	17°47′S	D-7,J-13[3],K-14	J-13(30)	DQ420392	DQ420584	DQ420456	DQ420520
40	Alice Springs [Il Parpa Swamp][c]	134°26′E	24°10′S	B-2[3],?-22	B-2(45)	DQ420368	DQ420560	DQ420432	DQ420496
					B-2(96)	DQ420370	DQ420562	DQ420434	DQ420498
					22(257)	DQ420411	DQ420603	DQ420475	DQ420539
41	Berry Springs[c]	130°58′E	12°42′S	K-14[3],N-17	N-17(91)	DQ420404	DQ420596	DQ420468	DQ420532
42	Jim Jim Ck[d]	133°05′E	13°05′S	K-14[3],N-17,	K-14(63)	DQ420396	DQ420588	DQ420460	DQ420524
				N-18[2]	N-17(253)	DQ420405	DQ420597	DQ420469	DQ420533
					N-18(93)	DQ420412	DQ420604	DQ420476	DQ420540
					N-18(94)	DQ420413	DQ420605	DQ420477	DQ420541
43	Mataranka[b, c]	133°04′E	14°56′S	J-13					
44	Ormiston G.[b]	132°43′E	23°37′S	B-2[3],?-3	B-2(58)	DQ420369	DQ420561	DQ420433	DQ420497
					3(224)	DQ420423	DQ420615	DQ420487	DQ420551
45	Avon R. Shire[Woodpile][c]	147°23′E	38°03′S	C-4,C-6	C-6(199)	DQ420417	DQ420609	DQ420481	DQ420545
46	Echuca	144°45′E	36°08′S	A-1[4],G-10	G-10(66)	DQ420383	DQ420575	DQ420447	DQ420511
47	Gunbower[c]	144°22′E	35°58′S	A-1[4]	A-1(22)	DQ420365	DQ420557	DQ420429	DQ420493
48	Holland landing[c]	147°28′E	38°04′S	C-4,C-5					
49	Kerrang Shire[c]	143°55′E	35°44′S	A-1[5]					
50	Marley Pt[c]	147°15′E	38°05′S	C-4[3]					
51	Meerlieu[c]	147°23′E	38°01′S	C-4[2],C-6	C-6(251)	DQ420419	DQ420611	DQ420483	DQ420547
52	Mildura[b, c]	142°10′E	34°11′S	A-1[2]					
53	Devils Ck[b]	148°15′E	41°30′S	E-8[5]	E-8(89)	DQ420377	DQ420569	DQ420441	DQ420505
54	Glencoe Swamp[b]	148°15′E	41°30′S	E-8[3],P-20	P-20(211)	DQ420415	DQ420607	DQ420479	DQ420543
55	Jordan R. [nr Jerrico Jnc]	147°11′E	42°28′S	E-8[2],P-20[2]	P-20(88)	DQ420414	DQ420606	DQ420478	DQ420542
56	Upper Turners Marsh[b]	147°13′E	41°26′S	E-8[5]					
57	Cobdogla	140°25′E	34°15′S	A-1					
58	PNG Highlands [27/98][b]	143°53′E	5°38′S	M-16[4]	M-16(111)	DQ420401	DQ420593	DQ420465	DQ420529
					M-16(112)	DQ420402	DQ420594	DQ420466	DQ420530
					M-16(114)	DQ420403	DQ420595	DQ420467	DQ420531
59	PNG lowlands [26/97][b]	147°06′E	9°25′S	L-15[4]	L-15(119)	DQ420399	DQ420591	DQ420463	DQ420527
60	PNG lowlands [41/97][b]	146°35′E	8°58′S	L-15[4]	L-15(115)	DQ420398	DQ420590	DQ420462	DQ420526
61	PNG lowlands [170/97][b]	147°38′E	9°57′S	L-15[4]	L-15(124)	DQ420400	DQ420592	DQ420464	DQ420528

Species designation shows the species (A–Q) followed by the ITS2 variant (1–23), with the sample size shown in square parentheses if greater than one. Question marks indicate that a decision concerning species status was not made. GenBank accession numbers are given for partial sequences of four genes (ITS2, EF-1α, COI and COII) for 64 of these specimens.

[a] Locations: 1–18 (Queensland), 19–35 (New South Wales), 36–39 (Western Australia), 40–44 (Northern Territory), 45–52 (Victoria), 53–56 (Tasmania), 57 (South Australia), 58–61 (Papua New Guinea).

[b] Specimens collected as larvae.

[c] Specimens collected by carbon dioxide-baited light trap.

[d] Specimens collected by night landing catches.

stored at room temperature for 3–5 min. Tubes were centrifuged at 4000*g* for 2 min to pellet the DNA and the supernatant was removed by decanting. DNA was washed twice, each time by addition of 0.6 ml of 75% ethanol, mixing by inverting the tube several times, followed by centrifugation at 4000*g* for 2 min and removal of the supernatant by decanting. Tubes were air-dried then DNA solubilized by addition of 30 μl of 8 mM NaOH. The DNA solution was adjusted to pH 8 by addition of 3 μl of 0.1 M HEPES.

2.3. Markers and primers used in this study

The selection of molecular markers included nDNA and mtDNA as well as protein-coding and non-protein-coding gene regions. Primers were selected to obtain partial sequences for ITS2, exons of elongation factor-1 alpha (EF-1α), cytochrome *c* oxidase subunit I (COI), and COII. Internal primers were designed for this study based on conserved sequence areas of a subset of sequences from the Annulipes Complex or through alignments of mosquito sequences obtained through GenBank. The nDNA protein coding region EF-1α has been used in numerous phylogenetic studies (e.g. Regier et al., 2000) including in combination with mtDNA genes (Monteiro and Pierce, 2001), and was advocated as a standard molecular marker in insect molecular systematics by Caterino et al. (2000). EF-1α occurs as two copies in some Diptera, such as on chromosomes 2 and 3 in *Drosophila melanogaster* L., one copy of which contains multiple introns (Danforth and Ji, 1998). Complete genome sequencing of *Anopheles gambiae* Giles reveals a single copy of EF-1α located on chromosome 2 (GenBank). In the absence of contradictory information and for the purposes of this study we assumed a single copy in other anophelines including the Annulipes Complex. Nevertheless, to protect ourselves from the unlikely possibility of paralogy, we designed the leading edges of the *An. annulipes* primers with an abundance of mismatches to the EF-1α copy on chromosome 3 of *D. melanogaster*. Our primers EF-1α F (5′-tgatttcatcaagaacatgatcac-3′) and EF-1 α R3 (5′-ggtggt-

tcagcacgatgacctg-3′) were developed specifically for the Annulipes Complex from the alignment of 850 bp EF-1α sequences from more than 20 *Anopheles* species including *An. annulipes* (Wilkerson et al. unpublished and GenBank). Amplifications with these primers resulted in a single intronless amplicon of approximately 750 bp. The rapidly evolving, non-coding ITS2 of rDNA has been found extremely useful for discriminating closely related insect species, including anopheline mosquitoes (e.g. Beebe and Saul, 1995; Mukabayire et al., 1999; Xu et al., 1998). For ITS2, the primers ITS2 F.3 (5′-tccggtgtgtgtaggattga-3′) and ITS2 B (5′-tatgcttaaattc aggggt-3′, Beebe and Saul, 1995) resulted in a product of approximately 300 bp. Cytochrome oxidase subunit I possesses a mix of highly conserved and variable regions (Lunt et al., 1996). We used the amplicon from primer sites UEA9 (5′-gtaaacctaacatttttcctcaaca-3′) to UEA10 (5′-tccaatgcacta atctgccatatta-3′) of COI (approximately 300 bp) following Lunt et al. (1996) and Zhang and Hewitt (1996), who suggested that this gene region was suited for population genetic or sibling species studies due to its fast evolution. The mtDNA gene encoding COII showed promise as a good phylogenetic marker for four species of the Annulipes Complex (Foley et al., 1998) and the primers COII A3 (5′-catcaatgat attgaag-3′, Foley et al., 1998) and Bt-LYS (5′-gtttaagagaccag tacttg-3′, Liu and Beckenback, 1992) resulted in approximately 400 bp of product.

2.4. PCR, sequencing and alignment

For ITS2, each 50 μl PCR reaction contained 10 mM Tris–HCl (pH 8.3) and 50 mM KCl (as 10X PCR Buffer, Applied Biosystems, Foster City, CA), 0.2 mM of each dNTP, 1.0 U AmpliTaq DNA polymerase (Applied Biosystems, Foster City, CA), 2.0 mM $MgCl_2$, 0.2 μM of each primer, DMSO, 10% BSA, and approximately 0.12–2.5 ng DNA template (1/33 of DNA from an adult mosquito abdomen). The reaction mix for EF-1α was as for ITS2 except that DMSO was omitted. For COI and COII, each 50 μl PCR reaction contained the same concentrations of ingredients as for ITS2 and EF-1α except that DMSO and BSA were omitted and the concentration of $MgCl_2$ was 3.0 mM. A Perkin Elmer Cetas 480 thermal cycler was used. The PCR program for ITS2 involved: 3 cycles at 94 °C for 1:00 min, 37 °C for 1:30 min, and 72 °C for 1:40 min; 45 cycles at 94 °C for 0:40 min, 48°C for 0:30 min, and 72 °C for 0:50 min; and 72 °C for 7:00 min. For EF-1α the program was: 94 °C for 3:00 min; 45 cycles at 94 °C for 0:40 min, 54 °C for 0:30 min, and 72 °C for 1:00 min; and 72 °C for 7:00 min. The program for COI and COII was: 94 °C for 2:00 min; 5 cycles at 94 °C for 1:00 min, 37 °C for 0:40 min, and 72 °C for 0:40 min; 50 cycles at 94 °C for 0:45 min, 48 °C for 0:30 min, and 72 °C for 1:00 min; and 72 °C for 7:00 min. For both ITS2 and EF-1α, preheating the block to 94 °C improved amplification. PCR product was separated on a 2% agarose gel and visualized with ethidium bromide stain. Fragment sizes were estimated by comparison with molecular weight standards.

PCR products were purified using PEG precipitation (20% polyethylene glycol 8000/2.5 M NaCl) and sequenced directly using the primers listed above. Sequencing reactions were carried out on both strands of DNA using ABI Big Dye chemistry (PE Applied Biosystems), and the sequences were generated with an ABI 3100 automated sequencer. Complimentary strands were combined into consensus sequences and questionable base calls were resolved using Sequencher 4.2. For ITS2, only sequences of *An. annulipes s.l.* were aligned, with CLUSTALX (Thompson et al., 1997), due to the difficulty of aligning ingroup and outgroup taxa to each other. The gap opening cost and the gap extension cost were set to the default (i.e. 15 and 6.66, respectively), and subsequent adjustments to the automated alignment were done manually. Numbers of variable and parsimony informative sites, and proportion of nucleotide differences were calculated by MEGA 3.0 (Kumar et al., 2004).

2.5. Phylogenetic analysis

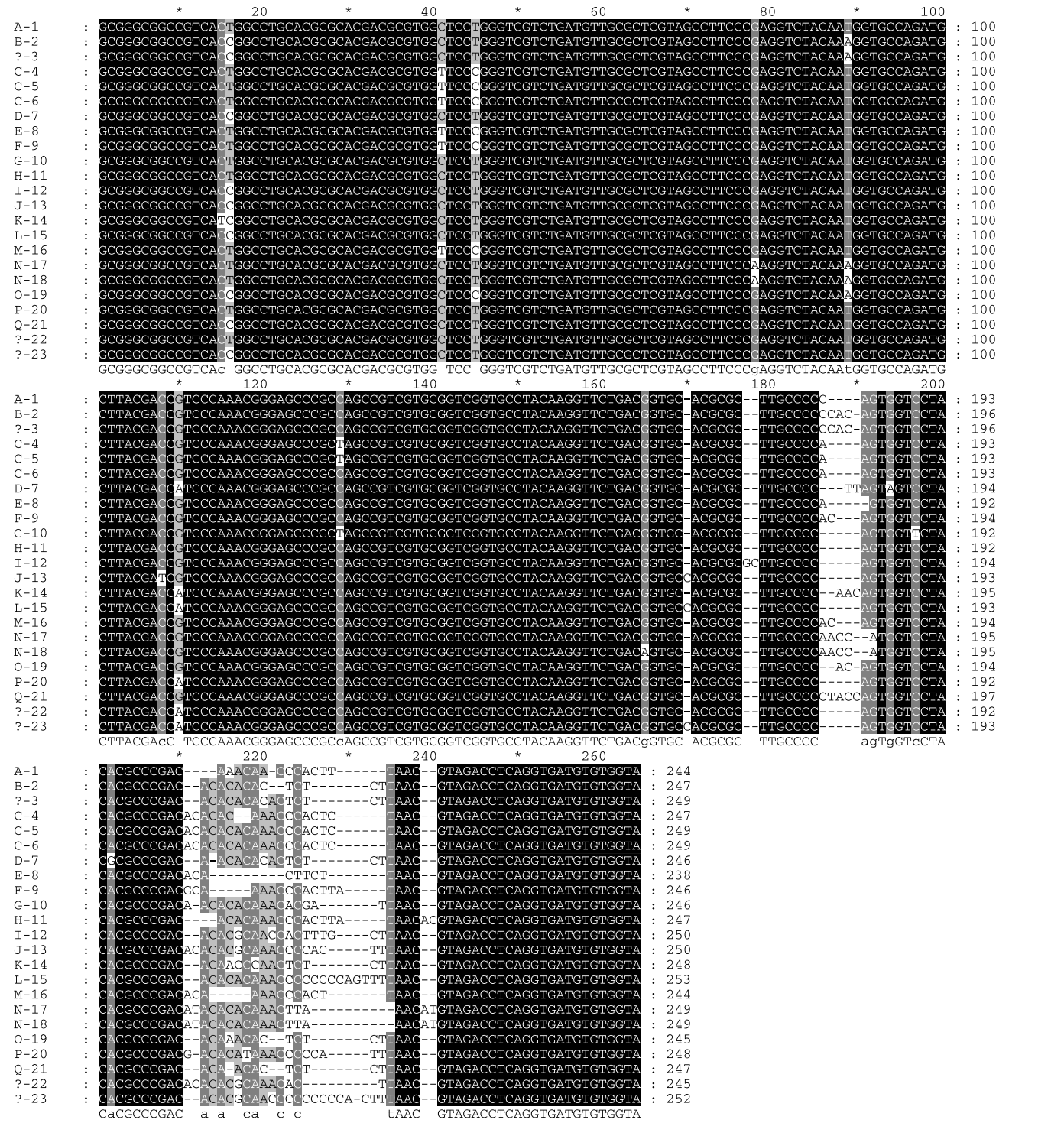
Preliminary analysis indicated that ITS2 sequence was diagnostic for *An. annulipes* species A–G, therefore, up to three individuals of each ITS2 variant from disparate localities were sequenced for COI, COII and EF-1α for inclusion in the phylogenetic study.

Phylogenetic reconstruction was performed using Bayesian analysis (Markov Chain Monte Carlo, MCMC) with MrBayes 3.0b4 (Huelsenbeck and Ronquist, 2001) and maximum parsimony (MP) analysis with PAUP* version 4.0b10 (Swofford, 2002) using *An. farauti* as outgroup. Maximum parsimony and Bayesian analyses were made on the combined concatenated data set of all four genes, then rDNA and mtDNA data sets were analysed separately by the Bayesian method.

For the combined Bayesian analysis, the models of substitution for each of 10 partitions, comprising genes and codon positions, were derived from MrModeltest 2.2 (Nylander, 2004). The Akaike information criterion rather than the hierarchical likelihood ratio test was used for model selection following Posada and Buckley (2004). Within MrBayes, three heated chains and a single cold chain were used in all MCMC analyses, which were run for 10,000,000 generations, sampling 1 tree every 1000 generations; trees obtained before convergent and stable likelihood values were discarded (i.e. a 2500 tree 'burn-in'). Four independent runs, each started from different, randomly chosen trees, were performed to assess convergence. Bayesian analyses of mtDNA and nDNA used the same settings but only those gene and codon partitions appropriate for the data set (i.e. six and four partitions, respectively). Majority rule consensus phylograms and posterior probabilities for nodes were assembled from all post burn-in trees (i.e. 30,004 trees per analysis).

Most parsimonious trees were determined with the heuristic search algorithm (100 random stepwise addition replicates) using the TBR swapping algorithm and unordered, equally weighted characters, with the MULTREES option

Table 2
Sequence of 23 ITS2 variants from specimens of *Anopheles annulipes s.l*



Sequences are labelled to indicate the species (letters A–Q) and the ITS2 variant (Nos. 1–23). Nucleotide position 1–241 of the aligned sequence is the 3′ end of ITS2 and 242–264 is the 5′ end of 28S of rRNA. Question marks indicate that a decision concerning species status was not made. ∗ indicates the position of every alternate tenth basepair.

on (Swofford, 2002). Phylogenetic inference was conducted with the concatenated data including all four genes with gaps considered as missing bases. We assessed support for clades using 1000 bootstrap pseudoreplicates with 5 random addition sequence replicates per bootstrap, setting a maximum tree limit of 100 for each random addition sequence replicate and holding 10 trees at each step during stepwise addition.

The phylogenetic species concept of Wheeler and Platnick (2000) was used as a criterion for the recognition of species in this study. This concept states "a species is the

smallest aggregation of (sexual) populations or (asexual) lineages diagnosable by a unique combination of character states." However, we also considered ecological, chromosomal, cross-mating and allozyme evidence from previous studies, especially where this indicated assortative mating in nature.

2.6. Morphology

Adult female *An. annulipes* s.l. from throughout the range of the taxon, were scored for 14 morphological characters (Table 5) that have been identified by previous authors as variable (e.g. Lee and Woodhill, 1944), or potentially informative (Green unpublished; Liehne, 1991; Skuse, 1889; Taylor, 1943). Voucher specimens were viewed in alcohol under a dissecting microscope and a graticule eyepiece was used to measure lengths. As specimens were primarily kept for DNA analysis rather than morphology, most specimens were damaged, so not all specimens could be scored for all characters.

3. Results

3.1. Phylogenetics and distribution

Partial sequences were obtained for ITS2 and 28S (240–264 bp). Twenty-three *An. annulipes* s.l. ITS2 variants (Table 2) were found from direct sequencing of 243 specimens from 61 sites (Table 1). The 3′ region of *An. annulipes* s.l. ITS2 contains a hypervariable region spanning 24 bp in our alignment that explains most of the differences between these ITS2 variants. Representatives of each of these ITS2 variants were chosen for further sequencing and analysis ($n = 59$). GenBank accession numbers are given in Table 1 for partial sequences of EF-1α (624 bp), COI (261 bp), and COII (384 bp), and Table 3 shows a breakdown of constant, variable and parsimony informative sites for these sequences. *Anopheles annulipes* s.l. mtDNA sequences were 603 bp long and nDNA sequences were 888 bp (including indels), giving a total of 1491 bp for phylogenetic analysis. The best models of DNA evolution for the 10 data partitions predicted by MrModeltest for the Likelihood Ratio Test and the Akaike information criterion are shown in Table 4. Fig. 1 shows the tree from the partitioned Bayesian Analysis of all four genes and Figs. 2 and 3 shows the results for mtDNA and nDNA, respectively. We found 5247 most parsimonious trees of length 603 steps (CI = 0.538, RI = 0.836, RC = 0.450, HI = 0.462) and the strict consensus tree for the MP analysis is shown in Fig. 4 along with Bootstrap values greater than 50%.

The results (Figs. 1–4) show that the Annulipes Complex is monophyletic with strong branch support. The nDNA phylogeny (Fig. 3) had only 59% posterior probability support for the ingroup, however, this analysis is affected by the lack of ITS2 data for outgroup taxa. Two major clades within the Annulipes Complex were recovered (Fig. 1). Clade 1 has a northern (tropical) distribution in Australia whereas clade 2 is mainly southern (temperate) with a zone of overlap in QLD (see map, Fig. 1). The PNG specimens (115, 119, 124) from the tropical lowland sites clustered with clade 1 and the specimens (111, 112, 114) from the more temperate highland location clustered with clade 2 (e.g. Fig. 1), suggesting that climate rather than latitude is the prime indicator of distribution.

The previously recognized species A–G are monophyletic and exhibit single unique ITS2 sequences, suggesting that ITS2 will be diagnostic for other species within this complex. Considering the relationships derived from sequences of four genes (Figs. 1–4), we recognize 17 species-level clades within the Annulipes Complex. We decided to continue the species-lettering scheme of Green (unpublished) and Booth and Bryan (1986) as formal taxonomic descriptions and naming of species has yet to be done. We therefore designate members of these 17 clades as *An. annulipes* species A–Q. Each species possesses a different ITS2 sequence.

Anopheles annulipes species C and N each comprise more than one ITS2 variant but these intraspecific variants exhibit minor (<3 bp) sequence differences, are paraphyletic in most reconstructions, and are likely to be the result of incomplete homogenisation of multiple copies of ITS2. Three specimens (22, 23 and 224) with unique ITS2 sequences have a sister-group relationship to one of the 17 species in most analyses, and are possibly additional

Table 3
Sequence data description of 59 specimens of *Anopheles annulipes* s.l. used in the phylogenetic analysis

Gene	Partition	Length	Constant	Variable	Parsimony informative
COI	Codon1	86	81	5	5
COI	Codon2	86	85	1	1
COI	Codon3	86	53	33	30
COII	Codon1	115	108	7	5
COII	Codon2	115	113	2	1
COII	Codon3	115	69	46	41
EF-1α	Codon1	208	204	4	3
EF-1α	Codon2	208	208	0	0
EF-1α	Codon3	208	163	45	34
ITS2		264	226	38	36
ITS2	Less gapped	230	213	17	16

Table 4
Best models predicted by MrModeltest 2.2 for outgroups plus 59 specimens of *Anopheles annulipes* s.l. used in the MrBayes phylogenetic analysis

Gene	Partition	hLRT	AIC
COI	Codon1	K80 + G	HKY + I
COI	Codon2	JC	F81
COI	Codon3	GTR + G	HKY + G
COII	Codon1	SYM	GTR + G
COII	Codon2	F81	F81
COII	Codon3	HKY + G	GTR + G
EF-1α	Codon1	HKY	HKY + I
EF-1α	Codon2	F81	F81
EF-1α	Codon3	GTR + G	GTR + G
ITS2	ITS2	GTR + G	GTR + I

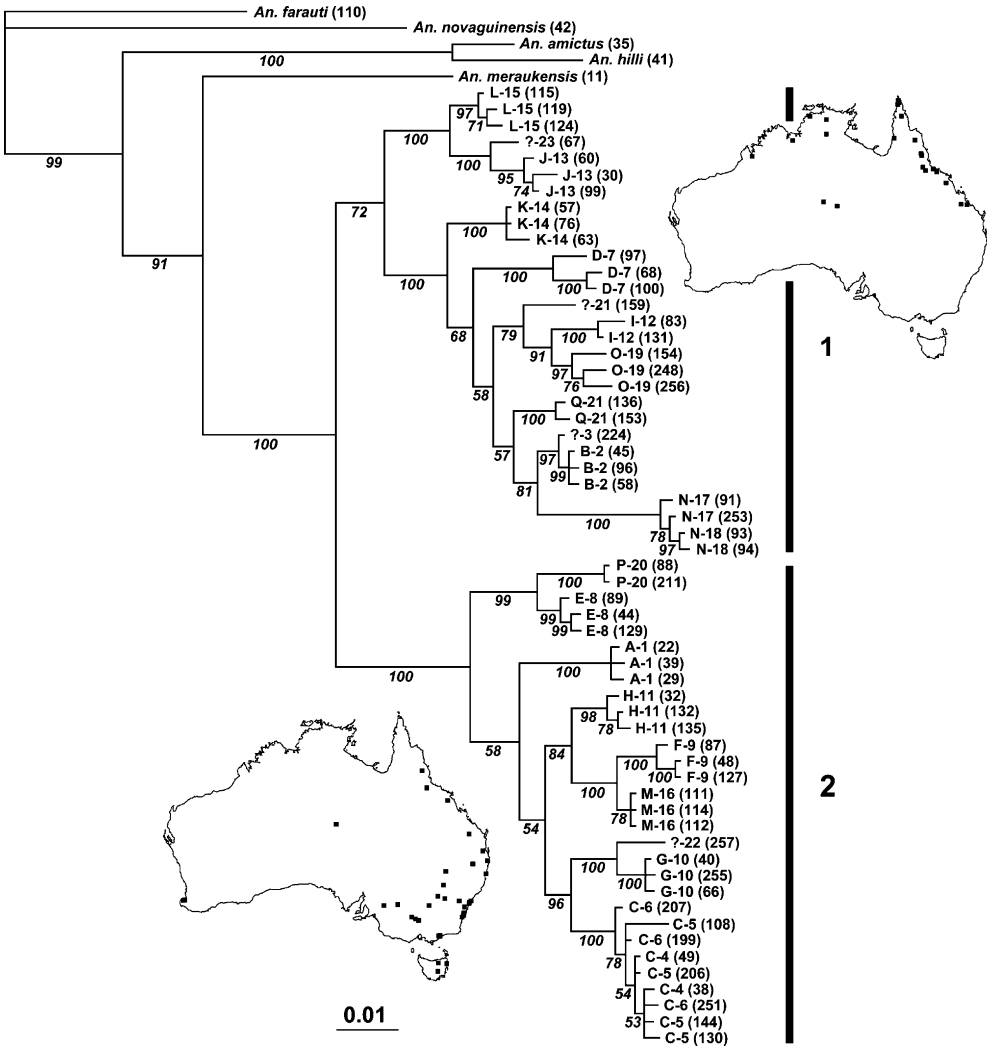


Fig. 1. Bayesian phylogeny of four genes (COI, COII, EF-1α and ITS2) from *Anopheles annulipes* s.l. and five other species of *Anopheles* (*Cellia*) from Australia (with *An. farauti* as outgroup). The majority rule phylogram and posterior probabilities resulting from 30004 post burn-in trees from a partitioned Bayesian analysis are shown, along with clades 1 and 2 and the distribution of sample points in these clades (PNG and Lord Howe Is. not shown). Specimens of *An. annulipes* s.l. are labelled to indicate the species (letters A–Q), the ITS2 variant (Nos. 1–23), and the specimen identification code (in parentheses). Question marks indicate that a decision concerning species status was not made. Bar = expected changes per site.

species. Further specimens are needed to verify the reciprocal monophyly of these ITS2 variants. Specimen 224 (ITS2 variant 3) is phylogenetically, geographically (i.e. sympatric at Ormiston Gorge, NT), and morphologically (i.e. possessing a "gingery" colored integument) closest to species B. ITS2 variant 22 is most closely related to species G but is derived from the specimen labelled species D in Foley et al. (1998). The identification of this specimen by these authors appears to be incorrect. Specimens matching the distribution (i.e. northern WA) and morphology (small, dark with reduced black on palpus III) of chromosomally identified species D (Green, unpublished; Liehne, 1991) were found in the present study, and are designated species D. ITS2 variant 23(67) from the Townsville region is most closely related to species J, but these differ substantially in their ITS2 sequences, and Townsville has a different climate than the locations where species J was found (data not shown), suggesting that they may be a different species. Specimen 159 (from Dilulu, QLD) possessed the same ITS2 sequence as other members of species Q (from Basalt River) and clustered with these in the nDNA Bayesian analysis (Fig. 3) and MP analysis (Fig. 4), suggesting that it is species Q. However, this specimen did not cluster with species Q in the 4-gene Bayesian analysis (Fig. 1) due to its anomalous mtDNA sequence that appeared most closely related to species O (see Fig. 2). Finally, the type 1 NT form of Booth and Bryan (1986) from Mataranka is an additional species that we have designated as species K.

For the mtDNA data (603 bp), the overall mean distance between species was 4% (range 0.3–6.8%) and the mean within group distance was 0.4% (range 0.0–2.0%). The smallest between-group distance was between species G and ITS2 variant 22 and the largest within-group distance was for species Q.

The composition of the Annulipes Complex in each geographic area based on DNA sequence data from the present

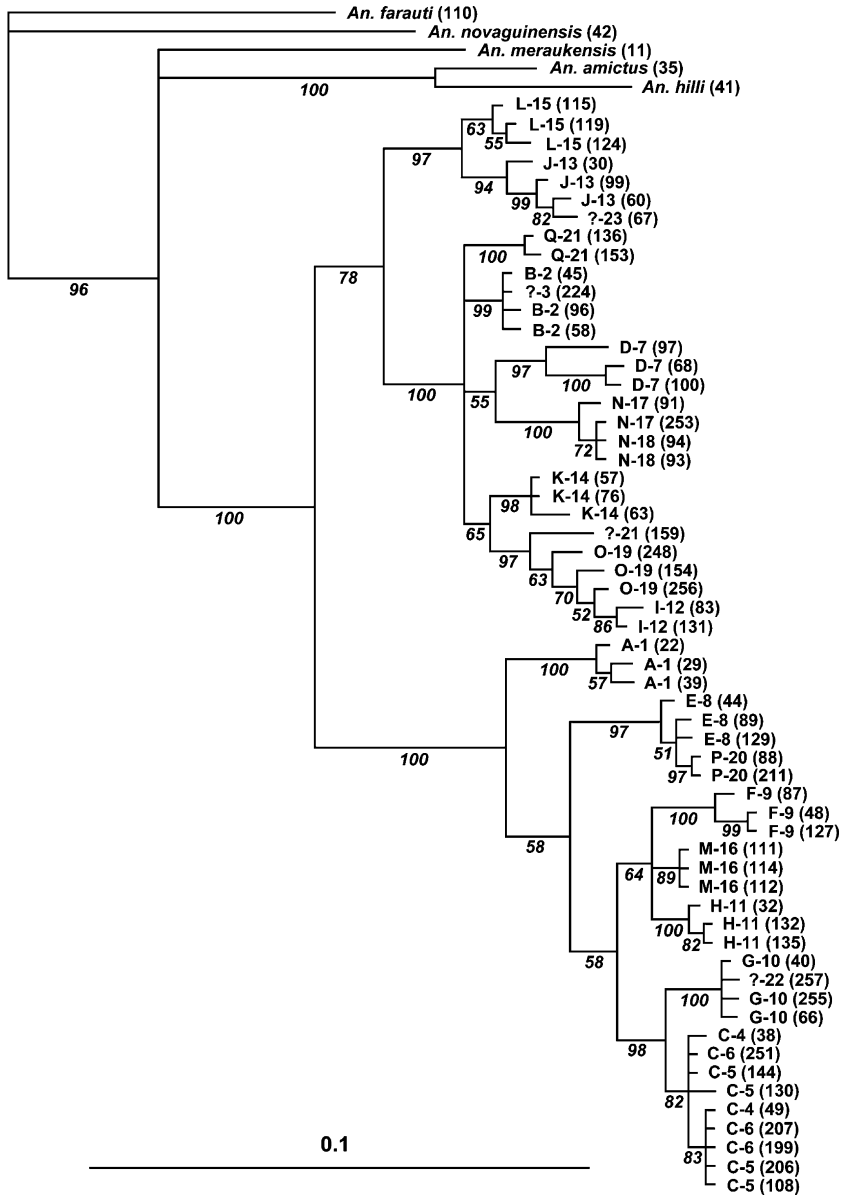


Fig. 2. Bayesian phylogeny of mtDNA genes (COI and COII) from *Anopheles annulipes* s.l. and five other species of *Anopheles* (*Cellia*) from Australia (with *An. farauti* as outgroup). The majority rule phylogram and posterior probabilities resulting from 30004 post burn-in trees from a partitioned Bayesian analysis are shown. Specimens of *An. annulipes* s.l. are labelled to indicate the species (letters A–Q), the ITS2 variant (Nos. 1–23), and the specimen identification code (in parentheses). Question marks indicate that a decision concerning species status was not made. Bar = expected changes per site.

study was: Victoria (VIC) (species A, C, G); NSW (species A, C, E, F, G, H); QLD (species D, E, F, H, I, J, K, O, Q); NT (species B, J, K, N); WA (species A, D, J, K); TAS (species E, P); South Australia (species A); and PNG (species L, M). In addition, Booth et al. (1987) recorded polytene chromosome-identified species B in WA, species A in TAS, and species C in QLD.

3.2. Morphology

The scores for the morphological characters are given in Table 5. Proboscis coloration was divided into three categories, all-dark, distinctly half pale, and intermediate. Intermediate coloration included specimens with patches of pale scales and specimens with a scattering of dark scales within patches of pale scales. The numbers of specimens are small and this limits conclusions about species differences. Species A is polymorphic for proboscis coloration and was the only species containing members with an all-dark proboscis. Species A specimens with an all-dark proboscis occurred sympatrically with specimens with a partially pale proboscis, for instance at Gunbower, VIC. Other species had variable amounts of pale on the proboscis and further sampling is needed to confirm that they do not exhibit the all-dark condition. Interspecific variation in the proportion of the palpomere that is dark scaled may be diagnostic for some species. Others (Green, unpublished; Liehne, 1991) suggested that species D could be separated from species A and B based on this character. Our results support this conclusion, but species D could be confused with species H, L and M on this character. Species B and ITS2 variant 3 possessed a

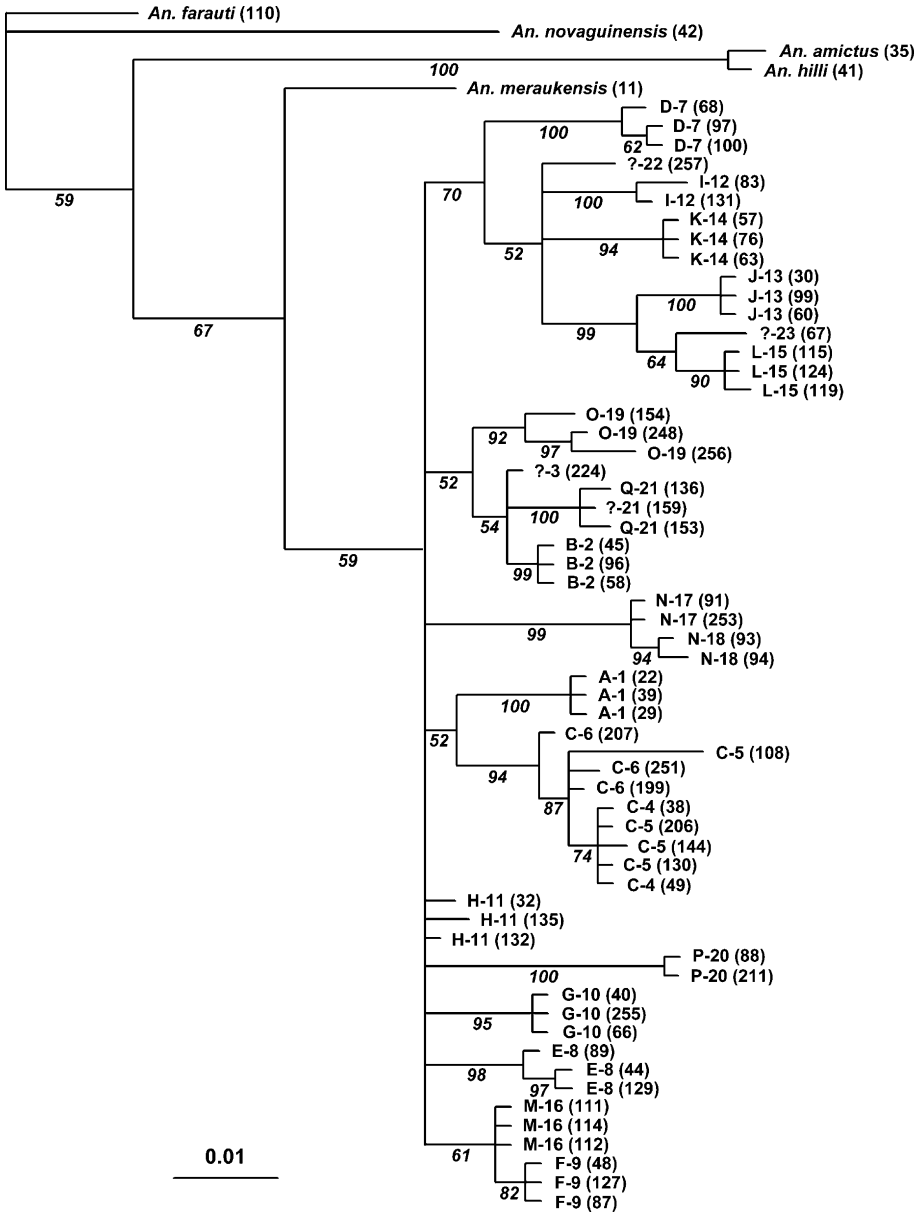


Fig. 3. Bayesian phylogeny of nDNA genes (EF-1α and ITS2) from *Anopheles annulipes* s.l. and five other species of *Anopheles* (*Cellia*) from Australia (with *An. farauti* as outgroup). The majority rule phylogram and posterior probabilities resulting from 30004 post burn-in trees from a partitioned Bayesian analysis are shown. Specimens of *An. annulipes* s.l. are labelled to indicate the species (letters A–Q), the ITS2 variant (Nos. 1–23), and the specimen identification code (in parentheses). Question marks indicate that a decision concerning species status was not made. Bar = expected changes per site.

"gingery"-colored integument (Green, unpublished; Liehne, 1991) that resulted in a low score for integument coloration in our data (see Table 5). A light (i.e. pale) colored thoracic integument was also recorded for species H and I. This contrasted with the generally dark species A, D, E and G. Species I and L had few upper pro-episternal setae (means: 2.3–2.8) and higher numbers occurred in species A, H and ITS2 variant 3 (means: 6.6–6.8). The number of dark spots on veins R and R_1 is used as a character to separate the Annulipes Complex (5–10 spots) from the Punctulatus Complex (9–17 spots), for instance in the key of Lee et al. (1987). One specimen of each of species I and N possessed 11 spots.

Generally members of clade 1 were lighter colored, less setose, and smaller (as measured by wing length) than clade 2 members. For instance, mean wing length was smaller for specimens of Clade 1 (3.326 mm ± 0.460, $n = 50$) than Clade 2 (4.098 mm ± 0.389, $n = 44$). How much of this variation is due to environment and how much to genetics is unknown. Bryan et al. (1991) found that various wing measurements, including wing length (tip to alula) were not significantly different between species A and G at Griffith, NSW, which agrees with the current findings for these species.

4. Discussion

This study demonstrates the monophyly of the Annulipes Complex and supports allozyme evidence (Foley et al., in press) for the existence of two major clades, one

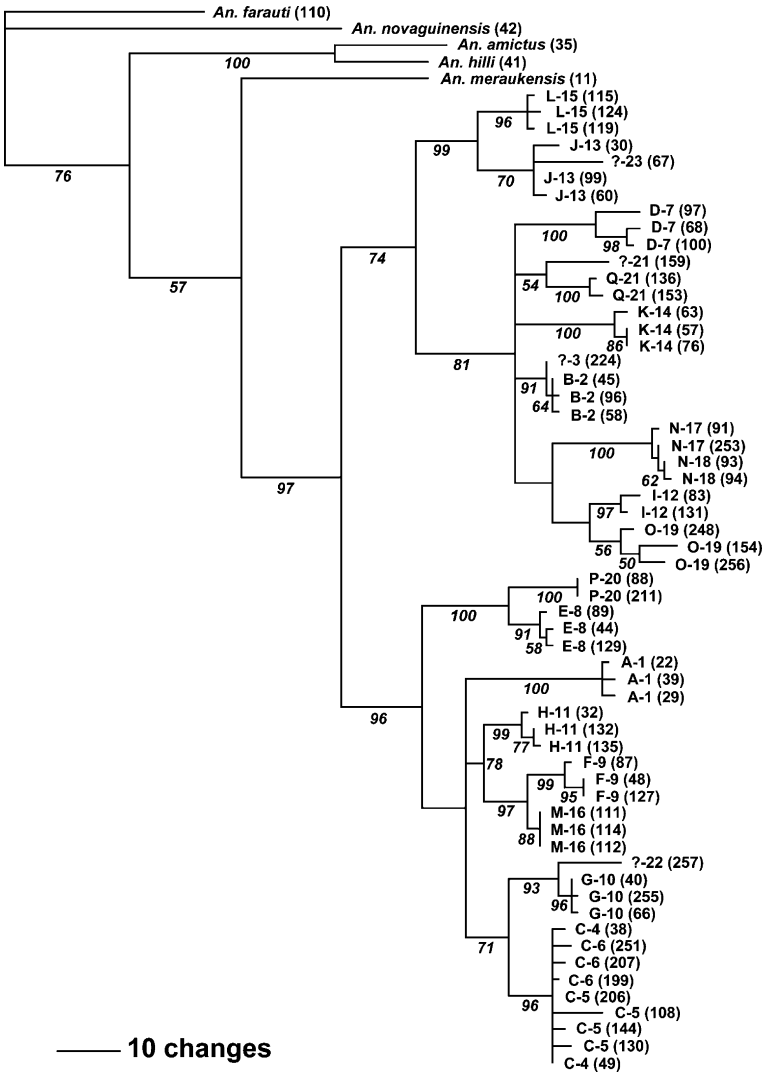


Fig. 4. Maximum parsimony phylogeny of four genes (COI, COII, EF-1α and ITS2) from *Anopheles annulipes* s.l. and five other species of *Anopheles* (*Cellia*) from Australia (with *An. farauti* as outgroup). The strict consensus phylogram of 5247 most parsimonious trees based on the equally weighted heuristic search is shown along with Bootstrap values greater than 50%. Specimens of *An. annulipes* s.l. are labelled to indicate the species (letters A–Q), the ITS2 variant (Nos. 1–23), and the specimen identification code (in parentheses). Question marks indicate that a decision concerning species status was not made. Bar = expected changes per site.

occurring in the north and one largely occurring in the south of Australia. Using molecular data we demonstrated the reciprocal monophyly of the seven chromosomally defined species A to G, and identified 10 additional species (species H–Q) and three possible new species (ITS2 variant 3, 22 and 23). Thus, *An. annulipes* s.l. is the most species-rich anopheline complex studied to date; by contrast the Gambiae Complex has 8 species (Harbach, 2004). The species complexity of *An. annulipes* s.l. may be the cause of much of the difficulty that taxonomists have had in explaining the extensive morphological variation within this taxon.

Although, ITS2 was diagnostic for the previously recognized species A–G, species C contained three sequence variants, sometimes occurring at the same collection site (e.g. Lord Howe Is). Multiple copies of rDNA units tend to become homogenized by a process called concerted evolution, where mutations rapidly spread to all members of the gene family even if there are arrays located on different chromosomes (e.g. Dover, 1982). However, cloning of ITS2 in various organisms has revealed that, when mutation rates are higher than homogenization, multiple copies of this gene region can occur within a species and even within an individual, with implications for phylogenetic analyses (e.g. Harris and Crandall, 2000). Despite this possible complication, direct sequencing of ITS2 is one of the most useful methods for identifying isomorphic species of mosquitoes (e.g. Wilkerson et al., 2004). The general concordance of phylogenetic trees based on mtDNA and nDNA indicates that ITS2 is a useful species marker within the Annulipes Complex.

Lee et al. (1987, pp. 136–138) gives a history of the taxonomic decisions concerning *An. annulipes.* Walker described this taxon as having the proboscis coloration "*proboscis ex parte testacea*" (i.e. part brownish), although

Table 5
Morphological character scores for 17 sibling species recognized within the Annulipes Complex of *Anopheles*

Character and state	Species																
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
Head. Proboscis. Apical pale scaling	(39)	(2)	(10)	(21)	(14)	(2)	(1)	(4)	(4)	(4)	(28)	(5)	(4)	(3)	(1)	(0)	(6)
Absent (0) Indistinct (1) Distinct (2)	012	12	12	12	12	2	1	12	2	12	2	2	2	2	2	—	12
Head. Palpomere III (length dark/total)	(7)	(2)	(10)	(10)	(9)	(3)	(2)	(4)	(3)	(3)	(9)	(5)	(4)	(2)	(1)	(1)	(4)
Mean	0.56	0.53	0.53	0.48	0.50	0.50	0.60	0.51	0.43	0.50	0.49	0.51	0.50	0.50	0.50	0.55	0.50
SD	0.05	0.04	0.04	0.04	0.00	0.00	0.00	0.03	0.06	0.00	0.03	0.02	0.00	0.00	—	—	0.00
Head. Palpomere IV (length dark/total)	(26)	(2)	(10)	(10)	(9)	(3)	(1)	(4)	(2)	(3)	(9)	(5)	(4)	(2)	(1)	(1)	(3)
Mean	0.45	0.42	0.30	0.19	0.22	0.25	0.40	0.22	0.16	0.28	0.22	0.20	0.19	0.29	0.25	0.40	0.30
SD	0.08	0.12	0.05	0.04	0.04	0.00	—	0.06	0.13	0.05	0.04	0.03	0.01	0.06	—	—	0.05
Head. Palpomere V (length dark/total)	(8)	(2)	(9)	(9)	(9)	(3)	(1)	(4)	(2)	(3)	(8)	(5)	(4)	(1)	(1)	(1)	(3)
Mean	0.61	0.50	0.41	0.27	0.42	0.36	0.60	0.27	0.38	0.36	0.30	0.28	0.29	0.50	0.50	0.67	0.41
SD	0.08	0.00	0.09	0.06	0.08	0.13	—	0.04	0.18	0.13	0.09	0.04	0.05	—	—	—	0.09
Thorax. Integument coloration[a]	(9)	(4)	(10)	(10)	(9)	(3)	(4)	(4)	(3)	(5)	(9)	(5)	(4)	(4)	(1)	(1)	(6)
Mean	1.4	0.0	1.3	1.4	1.5	1.2	1.4	0.4	0.5	1.2	1.1	1.0	1.1	1.0	1.0	1.0	0.8
SD	0.5	0.0	0.4	0.4	0.4	0.3	0.5	0.5	0.5	0.4	0.4	0.0	0.3	0.0	—	—	0.4
Thorax. Upper proepisternal setae[b]	(9)	(4)	(10)	(10)	(10)	(3)	(4)	(4)	(4)	(5)	(10)	(5)	(4)	(4)	(1)	(1)	(6)
Mean	6.6	4.3	6.1	5.2	6.0	6.0	6.3	6.8	2.3	5.8	4.9	2.8	5.0	3.5	6.0	6.0	3.5
SD	2.2	1.7	1.4	1.4	0.9	2.0	0.5	1.0	1.3	0.4	1.9	0.8	1.2	1.0	—	—	0.8
Thorax. Upper mesepimeral setae[b]	(9)	(4)	(10)	(10)	(10)	(3)	(4)	(4)	(4)	(5)	(10)	(5)	(4)	(4)	(1)	(1)	(6)
Mean	15.0	15.0	15.8	12.5	12.4	14.0	12.0	17.8	14.0	12.4	13.6	7.6	18.5	15.5	16.0	6.0	14.3
SD	2.7	2.0	2.7	2.2	2.1	2.0	1.6	7.0	1.6	4.6	2.7	1.7	4.4	2.5	—	—	4.3
Thorax. Upper meskatepisternal setae[b]	(10)	(4)	(10)	(9)	(9)	(3)	(4)	(4)	(4)	(5)	(10)	(5)	(4)	(4)	(1)	(1)	(6)
Mean	6.0	7.0	6.8	4.8	6.2	6.0	6.5	6.5	6.3	6.8	5.8	6.6	10.0	5.0	8.0	6.0	6.0
SD	2.3	3.8	2.5	1.9	1.9	2.0	1.9	1.9	1.7	2.3	1.5	1.7	3.3	1.2	—	—	1.8
Thorax. Lower meskatepisternal setae[b]	(9)	(4)	(10)	(10)	(9)	(3)	(4)	(4)	(4)	(5)	(10)	(5)	(4)	(4)	(1)	(1)	(6)
Mean	6.4	6.5	8.1	8.0	6.1	7.3	6.5	6.0	4.8	5.2	5.4	4.8	8.0	5.0	6.0	6.0	5.7
SD	0.9	1.9	2.0	2.7	0.3	2.3	1.9	1.6	1.5	1.8	1.0	1.1	1.6	1.2	—	—	1.5
Thorax. Post. mesepimeral scale patch	(9)	(2)	(10)	(10)	(10)	(3)	(3)	(4)	(4)	(5)	(10)	(5)	(4)	(4)	(0)	(1)	(6)
Absent (0) Present (1)	1	01	1	01	1	1	1	1	1	1	01	01	01	01	—	1	01
Wings. Fringe pale spots	(28)	(2)	(10)	(17)	(24)	(1)	(2)	(4)	(4)	(4)	(21)	(5)	(4)	(4)	(1)	(1)	(6)
Absent (0) Present (1)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Wings. Dark spots on veins R and R_1 (number)	(36)	(4)	(10)	(10)	(10)	(3)	(4)	(4)	(4)	(5)	(10)	(5)	(4)	(4)	(1)	(1)	(6)
Mean	6.8	6.3	7.8	5.8	6.8	6.7	7.5	9.0	6.0	6.8	6.1	7.4	8.8	6.5	6.0	5.0	6.5
SD	0.8	0.5	1.3	1.0	1.1	1.5	2.1	1.4	1.2	1.1	0.9	0.9	2.1	0.6	—	—	1.6
Wings. Veins M_1: R_3 (length)	(7)	(4)	(10)	(10)	(10)	(3)	(3)	(4)	(4)	(2)	(10)	(5)	(4)	(4)	(1)	(1)	(6)
Mean	0.67	0.69	0.72	0.72	0.72	0.71	0.77	0.70	0.68	0.73	0.69	0.69	0.74	0.72	0.68	0.75	0.71
SD	0.05	0.01	0.03	0.04	0.03	0.04	0.06	0.05	0.02	0.00	0.03	0.03	0.01	0.04	—	—	0.03
Wings. Length (mm)	(38)	(4)	(10)	(10)	(10)	(3)	(3)	(4)	(4)	(4)	(10)	(5)	(4)	(4)	(1)	(1)	(6)
Mean	4.28	4.01	4.07	3.36	4.04	3.98	4.25	3.94	3.60	3.38	3.16	2.84	4.24	3.19	3.38	3.75	3.35
SD	0.42	0.86	0.42	0.34	0.33	0.34	0.57	0.16	0.35	0.16	0.37	0.12	0.32	0.13	—	—	0.50

Numbers of specimens inspected for a character are given in brackets. SD = standard deviation.
[a] Light (0), medium (1), dark (2).
[b] Sum of both side.

the extent of dark coloration can no longer be determined as the specimen now has no head (*fide* P.F. Mattingly, *in litt*. 27.x.1970 in Lee et al., 1987; p. 138). According to Townsend (1990) "Although only one specimen is in the collection [of the Natural History Museum, London] Walker gave a range of measurements in his description, indicating that he had a syntypic series". Based on geographic distribution, the holotype would be a member of clade 2. However, Walker (1856, in Giles, 1900 p. 158) recorded the length of the wings of *An. annulipes* as 6–7 lines (=2.84–3.31 mm), which appears small for a clade 2 species (see Table 5). Inspection of the holotype by one of us (DF) indicates that wing vein ratio M_1:R_3 is 0.73 and that there are 5 spots on the R_1. Further sampling from TAS is needed before a decision about the identity of *An. annulipes* sensu stricto is made.

Foley et al. (in press) suggested species C and species E as candidates for *An. musivus*, an all-dark form with a clade 2 distribution, based on allozymes of specimens from locations nearby the type locality of Sydney (i.e. McCarrs Ck and Homebush). However, we also found species A at McCarrs Ck. According to Skuse (1889), *An. musivus* had a

wing length of 5.08 mm; only members of species A had comparably large wing lengths. These observations, and the observation that only species A was found with an all-dark proboscis, suggests that *An. musivus* is species A.

The identity of *An. mastersi* is less clear, although the geographic distribution indicates that it is a clade 2 species and the smaller wing size suggests that it is not species A. According to Taylor (1943), *An. perplexus* is found at locations in the NT and north QLD, suggesting that it is a member of clade 1. He reports that the ratio of the posterior to anterior forked cell is 3:4 (0.75), the wing length is 3.5 mm, and there are 7 spots on the R_1. Based on the distribution and morphology reported here, *An. perplexus* is most likely species J. Taylor (1943) described *An. persimilis* from Irvinebank, QLD but this description lacks detail. He found that females of this species lacked wing fringe spots but we did not observe this in any of our specimens. We found species F, K and O at Irvinebank, but the 5:8 ratio of posterior forked cell to anterior forked cell and the wing length of 3.5 mm given by Taylor (1943) suggest that this species was not species F. The female (but not the male) of *An. derricki* in Taylor (1943), collected from Irvinebank, is apparently *An. novaguinensis* (Lee and Woodhill, 1944). From the description by Taylor (1943), the male appears to be different from *An. persimilis* and species F, K and O must therefore be considered candidates.

According to an unpublished manuscript by Green, "species A and B can be separated easily on their coloration. Species B is gingery, and A is dark-grey to brown. Species D is very small, dark, and has a distinctive palpal ornamentation". Liehne (1991) describes species B as having the proboscis dark scaled, but our specimens of species B had a distinct apical pale area on the proboscis, suggesting that this trait is variable in this species. Green (unpublished) and Liehne (1991) states that species B has a gingery color to the scutal integument and scales, characteristics we found to be unique for species B and ITS2 variant 3.

Lee and Woodhill (1944) noted that the coloration of the proboscis in *An. annulipes* is variable and that the extremes can be found sympatrically (e.g. Seymour, VIC), as we found at Gunbower. Based on geographic distribution and the range of proboscis coloration reported in our study, the variation in morphology at Seymour could be explained solely by the presence of species A. Lee and Woodhill (1944) also note that the dark form is common in southern Australia, but is rare north of Sydney. This suggests that the dark form is only present in species of clade 2. Increased sampling throughout the range of *An. annulipes* is needed to establish the geographic range of the species within this complex.

The phylogenetic trees from MP and Bayesian analyses of combined data were largely concordant although some differences occurred. Specimen 159 has the same ITS2 sequence as species Q and clustered with this species in the four-gene MP analysis but not the Bayesian analysis. Separate Bayesian analysis of nDNA and mtDNA revealed that the mtDNA of this specimen does not cluster with that of other specimens of species Q. This pattern can occur due to incomplete lineage sorting of ancestral mtDNA or from introgressive hybridization. The extent and conditions under which introgression occurs between species within the Annulipes Complex is unknown but it is likely to be rare under normal conditions; of 6630 specimens of species A and G identified by allozymes at Griffith, NSW, less than 1.5% could not be identified, and no hybrid chromosomes were detected in the offspring of 103 wild-caught females from an area of sympatry (Foley and Bryan, 1991a). No evidence of recent hybridization could be detected from the nDNA in the present study.

Moritz and Cicero (2004) argue that the true test of the precision of DNA barcodes to assign individuals to species would include comparisons with sister species. The species rich Annulipes Complex appears to be well suited to such a test. Despite our COI sequence being much shorter (258 bp) than the 648 bp full barcoding sequence, we found that 65% (11) of the 17 sibling species recorded here (species A, C, D, F, G, H, J, K, L, M, and N) had unique COI sequences. This number of COI-diagnosable species increased to 13 if ITS2 variants 3, 22 and 23 are counted as separate species, the latter two having unique COI sequences. This observation suggests that DNA barcoding holds some promise for diagnosing species within the Annulipes Complex, and perhaps for other anophelines. However, our sequence does not overlap the barcoding sequence and may not vary in the same way. Also, our sample sizes were low and will need to be increased to determine the true extent of paraphyly, particularly for ITS2 variants 3, 22 and 23. Meyer and Paulay (2005) noted that incomplete lineage sorting of recently diverged species in taxonomically understudied groups is a potential problem for DNA barcoding (e.g. Hebert et al., 2003). Hebert et al. (2003) argued that a threshold of 3% COI sequence divergence would be sufficient to separate species. However, such a threshold would not separate all of the species of the Annulipes Complex. The threshold barcoding level for the Annulipes Complex, as for other organisms (e.g. Meyer and Paulay, 2005), may have to be set lower than 3% to minimize false negatives.

Based on an mtDNA molecular clock of 2.3% per million years (Brower, 1994) the combined COI and COII data (603 bp) indicates that *An. annulipes* s.l. and *An. meraukensis* diverged about 4 million years ago (mya), and that clades 1 and 2 diverged about 2.5 mya. These estimates place the divergence of clades 1 and 2 around the Plio-Pleistocene, during the start of the geologically recent period of repeated glaciations. This timing suggests that climate change was an early driver of species radiation, the phylogenetic signal of which is still evident in the divergent geographical distributions of clade 1 and 2 species. Species or ITS2 variants that diverged most recently (<100 kya) are species O and species I, ITS2 variant 22 and species G, and species E and species P. The endemic PNG species L and M may have arisen through allopatric speciation, due to the alternate separation and reconnection of Australia and NG. The sister species relationship of the PNG highland

species M and the Australian species F suggests that the ancestor of these species extended from Australia into NG, possibly through the cooler, higher elevation areas of the Great Dividing Range. Climate warming, and the separation of Australia and NG, may have restricted this species in PNG to highland areas (>1500 m, Cooper et al., in press), and in Australia to lower latitude and higher elevation tropical areas (e.g. Irvinebank; 950 m and Eungella; 650 m), allowing speciation to occur.

Anopheles annulipes s.l. was the presumptive vector of malaria in southeastern Australia in the past (Black, 1972) and the present study suggests that the vectors belong to clade 2. Candidate malaria vectors (i.e. that have been collected for this study biting humans, Table 1) include *An. annulipes* species A, C, G, K and N with the former three species belonging to clade 2. *Anopheles annulipes* s.l. has not been incriminated as a vector of malaria in PNG, which suggests that species L and M are medically unimportant. Foley et al. (in press) incriminated species A and E as vectors of myxomatosis based on distribution but more sampling is required to confirm this.

Many of the sampling locations in this study and the allozyme study of Foley et al. (in press) are the same, and in some cases samples came from the same pool of specimens. The number of clusters determined by allozymes and molecular data shows much agreement. Foley et al. (in press) suggested that 15–25 species occur in Australia based on Bayesian and tree-based clustering of allozyme data, and the present study supports the lower to middle part of this range. Conclusions in Foley et al. (in press) about the number of species within the different states of Australia (WA + NT, NSW + TAS + VIC, QLD) compares most closely with the results here for their Bayesian analysis for $K = 15$ groups and the tree-based approaches using the Nei's D measure of genetic divergence (which suggested 18 species in Australia). This comparison is closer still if the observations of species distribution in Booth et al. (1987) (see above) are taken into consideration. However, with further sampling, it is likely that additional species within the Annulipes Complex will be discovered.

The phylogenetic reconstructions here provide a framework for testing hypotheses about ecology and behavior of the sibling species of the Annulipes Complex. This complex is the most species-rich among anophelines and, as such, provides a unique mosquito model for investigations into the evolution of genes and traits of medical and veterinary importance.

Acknowledgments

Much of this research was performed while DHF held a US National Research Council Research Associateship Award at the Walter Reed Army Institute of Research. This project was not possible without specimens kindly supplied by the following: Tony Barnes, Ann-Marie Boyd, Peter Ebsworth, Ian Fanning, Chris Freebairn, Brian Kay, Con Lokkers, Melina Miles, Andrew Van Den Hurk, Gimme Walter (Queensland); David Booth, Stephen Doggett, Merilyn Geary, Richard Russell, Margaret Spencer (NSW); Brian Cordwell, Gary George, Michael Holland, M. Kennedy, Brian Watson (VIC); Tony Wright (WA); Tom Burkot, Jim Stott, Peter Whelan, (NT); Jo Kent, and Craig Williams (SA). We thank Cong Li for advice and assistance with molecular work. This research was performed under a Memorandum of Understanding between the Walter Reed Army Institute of Research and the Smithsonian Institution, with institutional support provided by both organizations. The material to be published reflects the views of the authors and should not be construed to represent those of the US Department of the Army or the Department of Defense. This article is published with the approval of the Director General Defence Health Services (Australia). The opinions expressed herein are those of the authors and do not necessarily reflect those of the Australian Defence Health Service or any extant health policy. Mention of a commercial product does not constitute an endorsement of the product by the Australian Defence Force.

References

Beebe, N.W., Saul, A., 1995. Discrimination of all members of the *Anopheles* Punctulatus Complex by polymerase chain reaction-restriction fragment length polymorphism analysis. Am. J. Trop. Med. Hyg. 53, 481–578.

Black, R.H., 1972. Malaria in Australia. Australian Government Publishing Service, Canberra.

Booth, D.R., Bryan, J.H., 1986. Cytogenetic and crossbreeding evidence for additional cryptic species in the *Anopheles annulipes* Walker complex (Diptera: Culicidae). J. Aust. Entomol. Soc. 25, 315–325.

Booth, D.R., Green, C.A., Bryan, J.H., 1987. The larval salivary gland polytene chromosomes of *Anopheles* (Cellia) annulipes s.l. Walker (Diptera: Culicidae). Aust. J. Zool. 35, 247–252.

Brandley, M.C., Schmitz, A., Reeder, T.W., 2005. Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scincid lizards. Syst. Biol. 54, 373–390.

Brower, A.V., 1994. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. Proc. Natl Acad. Sci. USA 91, 6491–6495.

Bryan, J.H., Foley, D.H., Geary, M., Carven, C.T.J., 1991. *Anopheles annulipes* Walker (Diptera: Culicidae) at Griffith, New South Wales. 3. Dispersal of two sibling species. J. Aust. Entomol. Soc. 30, 119–121.

Caterino, M.S., Cho, S., Sperling, F.A.H., 2000. The current state of insect molecular systematics: a thriving Tower of Babel. Annu. Rev. Ecol. Syst. 45, 1–54.

Cooper, R.D., Waterson, D.G.E., Frances, S.P., Beebe, N.W., Sweeney, A.W., 2006. The anopheline fauna of Papua New Guinea. J. Am. Mosq. Control Assoc. 22, 213–221.

Danforth, B.N., Ji, S., 1998. Elongation factor-1α occurs as two copies in bees: implications for phylogenetic analysis of EF-1 α sequences in insects. Mol. Biol. Evol. 15, 225–235.

Dover, G., 1982. Molecular drive—a cohesive mode of species evolution. Nature 299, 111–117.

Fenner, F., Ratcliffe, G.N., 1965. Myxomatosis. Cambridge University Press, Cambridge.

Foley, D.H., Bryan, J.H., 1991a. *Anopheles annulipes* Walker (Diptera: Culicidae) at Griffith. New South Wales. 1. Two sibling species in sympatry. J. Aust. Entomol. Soc. 30, 109–112.

Foley, D.H., Bryan, J.H., 1991b. *Anopheles annulipes* Walker (Diptera: Culicidae) at Griffith, New South Wales. 2. Biology and behaviour of two sibling species. J. Aust. Entomol. Soc. 30, 113–118.

Foley, D.H., Barnes, A., Bryan, J.H., 1992. *Anopheles annulipes* Walker (Diptera: Culicidae) at Griffith, New South Wales. 4. Phenology of two sibling species. J. Aust. Entomol. Soc. 31, 91–96.

Foley, D.H., Bryan, J.H., Yeates, D.K., Saul, A., 1998. Evolution and systematics of *Anopheles*: insights from a molecular phylogeny of Australasian mosquitoes. Mol. Phylogenet. Evol. 9, 262–275.

Foley, D.H., Bryan, J.H., Wilkerson, R.C. Species-richness of the *Anopheles* Annulipes Complex (Diptera: Culicidae) revealed by tree and model-based allozyme clustering analyses. Biol. J. Linn. Soc., in press.

Giles, G.M., 1900. A Handbook of the Gnats or Mosquitoes. John Bale, Sons & Danielsson, Ltd, London.

Green, C.A., 1972. The *Anopheles annulipes* complex of species. In: Abstracts 14th International Congress of Entomology, Canberra, p. 286.

Harbach, R.E., 2004. The classification of genus *Anopheles* (Diptera: Culicidae): a working hypothesis of phylogenetic relationships. Bull. Entomol. Res. 94, 537–553.

Harris, D.J., Crandall, K.A., 2000. Intragenomic variation within ITS1 and ITS2 of fresh water crayfishes (Decapoda: Cambaridae): Implications for phylogenetic and microsatellite studies. Mol. Biol. Evol. 17, 284–291.

Hebert, P.D.N., Cywinska, A., Ball, S.L., deWaard, J.R., 2003. Biological identifications through DNA barcodes. Proc. R. Soc. Lond. B Biol. Sci. 270, 313–322.

Huelsenbeck, J.P., Ronquist, F., 2001. MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.

Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief. Bioinform. 5, 150–163.

Lee, D.J., Woodhill, A.R., 1944. The Anopheline Mosquitoes of the Australasian Region. Monograph No. 2. University of Sydney, NSW.

Lee, D.J., Hicks, M.M., Griffiths, M., Debenham, M.L., Bryan, J.H., Russell, R.C., Geary, M., Marks, E.N., 1987. The Culicidae of the Australasian Region. vol. 5. Australian Government Publishing Service, Canberra.

Liehne, P.F.S., 1991. An Atlas of the Mosquitoes of Western Australia. Health Department of Western Australia, Perth.

Lin, C-P., Danforth, B.N., 2004. How do insect nuclear ribosomal genes compare to protein-coding genes in phylogenetic utility and nucleotide substitution patterns? Syst. Entomol. 30, 686–702.

Lunt, D.H., Zhang, D.X., Szymura, J.M., Hewitt, G.M., 1996. The insect cytochrome oxidase I gene: evolutionary patterns and conserved primers for phylogenetic studies. Insect Mol. Biol. 5, 153–165.

Meyer, C.P., Paulay, G., 2005. DNA Barcoding: error rates based on comprehensive sampling. PLoS Biol. 3, e422.

Monteiro, A., Pierce, N.E., 2001. Phylogeny of Bicyclus (Lepidoptera: Nymphalidae) inferred from COI, COII, and EF-1α gene sequences. Mol. Phylogenet. Evol. 18, 264–281.

Moritz, C., Cicero, C., 2004. DNA Barcoding: promise and pitfalls. PLoS Biol. 2, e354.

Mukabayire, O., Boccolini, D., Lochouran, D., Fontenille, D., Besansky, N.J., 1999. Mitochondrial and ribosomal internal transcribed spacer (ITS2) diversity of the African vector *Anopheles funestus*. Mol. Ecol. 8, 289–297.

Nylander, J.A.A., 2004. MrModeltest 2.2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Parer, I., Korn, T.J., 1989. Seasonal incidence of myxomatosis in New South Wales. Aust. Wildlife Res. 16, 563–568.

Paterson, H.E., James, S.H., 1973. Animal and plant speciation studies in Western Australia. J.R. Soc. West. Aust. 56, 31–36.

Posada, D., Buckley, T.R., 2004. Model selection and model averaging in phylogenetics: advantages of Akaike Information Criterion and Bayesian approaches over Likelihood Ratio Tests. Syst. Biol. 53, 793–808.

Regier, J.C., Mitter, C., Peigler, R.S., Friedlander, T.P., 2000. Phylogenetic relationships in Laiocampidae (Lepidoptera): initial evidence from elongation factor-1alpha sequences. Insect Syst. Evol. 31, 179–186.

Rohland, N., Siedel, H., Hofreiter, M., 2004. Nondestructive DNA extraction method for mitochondrial DNA analyses of museum specimens. BioTechniques 36, 814–821.

Russell, R.C., 1995. Arboviruses and their vectors in Australia: an update on the ecology and epidemiology of some mosquito-borne arboviruses. Rev. Med. Vet. Entomol. 83, 141–158.

Skuse, F.A.A., 1889. Diptera of Australia. Part V. – The Culicidae. Proc. Lin. Soc. NSW 3, 1717–1764.

Swofford, D.L., 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4.0b10. Sinauer Associates, Sunderland, MA.

Taylor, F.H., 1943. Contributions to a knowledge of Australian Culicidae. No. VI. Proc. Linn. Soc. NSW 68, 153–157.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTALX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucl. Acids Res. 25, 4876–4882.

Townsend, B.C., 1990. Culicidae. In: Townsend, B.C., Chainey, J.E., Crosskey, R.W., Pont, A.C., Lane, R.P., Boorman, J.P.T., Crouch, C.A. (Eds). A Catalogue of Types of Bloodsucking Flies. Occas. Pap. Nat. Hist. Mus. No. 7, pp. 35–152.

Wheeler, Q.D., Platnick, N.I., 2000. The phylogenetic species concept. In: Wheeler, Q.D., Meier, R. (Eds.), Species Concepts and Phylogenetic Theory: A Debate. Columbia Univ. Press, New York, pp. 55–69.

Wilkerson, R.C., Reinert, J.F., Li, C., 2004. Ribosomal DNA ITS2 sequences differentiate six species in the *Anopheles crucians* complex (Diptera: Culicidae). J. Med. Entomol. 41, 392–401.

Xu, X., Xu, J., Qu, F., 1998. A diagnostic polymerase chain reaction assay for species A and D of the *Anopheles dirus* (Diptera: Culicidae) species complex based on ribosomal DNA second internal transcribed spacer sequence. J. Am. Mosq. Control Assoc. 14, 385–389.

Zhang, D.X., Hewitt, G.M., 1996. Assessment of the universality and utility of a set of conserved mitochondrial COI primers in insect DNA. Insect Mol. Biol. 6, 143–150.